Regionalization versus Competition in Complex Cancer Surgery

by

Vivian Ho, Ph.D.
Department of Health Care Organization and Policy
The University of Alabama at Birmingham
Birmingham, AL

Martin Heslin, M.D.
Department of Surgery
The University of Alabama at Birmingham
Birmingham, AL

Please address correspondence to:

Vivian Ho
Department of Health Care Organization and Policy
University of Alabama Birmingham
RPHB 330
1665 University Blvd.
Birmingham, AL 35294-0022

TEL number: (205) 975-0532
FAX number: (205) 934-3347
E-mail address: vho@uab.edu

February, 2003

PRELIMINARY. DO NOT CITE.

*For presentation at the 14th Annual Health Economic Conference, RAND, April 27-28, 2003.

This research was supported by grant number RSGHP-03-076-01-PBP from the American Cancer Society. I am grateful to Hal Luft for helpful comments on an earlier draft of this paper.

ABSTRACT

This study analyzes the relation between procedure volume and outcomes, charges, and costs for Whipple surgery, which is a complex surgical procedure for pancreatic cancer patients. Both increased hospital volume and surgeon volume are associated with lower inpatient mortality rates and lower charges and costs per patient. Independent time effects are not associated with reduced mortality or lower costs. We find no evidence that increased local market concentration for the Whipple procedure influences inpatient mortality or average cost per patient. However, the results suggest that spillover learning from nearby hospitals may lead to lower charges per patient. For the Whipple procedure, the benefits of high procedure volume outweigh the benefits of hospital competition. Outcomes will improve, and charges and costs will decline with regionalization.

STRUCTURED ABSTRACT

Objective: To analyze the relation between procedure volume and outcomes, charges, and costs for Whipple surgery, which is a complex surgical procedure for pancreatic cancer patients.

Data Sources: State hospital discharge abstracts from Florida (1988 to 1999), New Jersey (1988 to 1998), and New York (1989 to 1999).

Study Design: Regression analyses were conducted to examine variations in inpatient mortality, patient charges, and patient costs as a function of annual Whipple procedure volume for the admitting hospital and the annual number of Whipple surgeries performed by the patient's surgeon. The regressions included a Herfindahl index of local market competition, year fixed effects, a proxy for local spillover learning, and control variables for patient casemix.

Data Collection: We extracted data for all 8,145 patients with an ICD-9-CM procedure code of 52.7 (radical pancreaticoduodenectomy) from the state hospital discharge abstracts.

Principal Findings: Both increased hospital volume and surgeon volume are associated with lower inpatient mortality rates and lower charges and costs per patient. Independent time effects are not associated with reduced mortality or lower costs. We find no evidence that increased local market concentration for the Whipple procedure influences inpatient mortality or average cost per patient. However, the results suggest that spillover learning from nearby hospitals may lead to lower charges per patient.

Conclusions: For the Whipple procedure, the benefits of high procedure volume outweigh the benefits of hospital competition. Outcomes will improve, and charges and costs will decline with regionalization.

Key Words: Volume, Outcomes, Regionalization, Hospital Competition, Pancreaticuduodenectomy.

INTRODUCTION

Past studies have determined that hospitals performing a higher number of surgical procedures tend to have better outcomes and lower costs. This relation has been identified for a number of procedures, including coronary artery bypass graft surgery, hip fracture surgery, coronary angioplasty, and eight major types of cancer resection (H. S. Luft et al. 1979;R. G. Hughes et al. 1988;E. L. Hannan et al. 1991;K. A. Phillips et al. 1995;J. D. Birkmeyer et al. 2002). These findings have led several researchers and policy makers to recommend regionalization of complex surgeries. Proponents of regionalization argue that concentrating surgery at a few high-volume facilities and eliminating low-volume providers will lead to improved patient outcomes and lower costs.

However, another branch of the literature finds that competition among hospitals can lead to improved outcomes and/or lower patient costs (D. P. Kessler, M. B. McClellan 2000;G. Gowrisankaran, R. Town 2002). Moreover, technological progress can also lead to lower mortality rates and lower costs per patient. Thus, hospitals of all sizes may improve over time regardless of their patient volume. In addition, regionalization could stifle the potential benefits of competition among several existing low-volume providers.

This manuscript uses the Whipple procedure as a case study to examine the potential benefits of regionalization versus competition in complex cancer surgery. The data for this study come from patient-level hospital discharge abstracts for Florida (1988-99), New Jersey (1988-98), and New York (1989-99). The dataset contains information on all patients (8,145) who received the Whipple procedure in these states and years. The two dependent variables of interest are inpatient mortality and patient-level charges. Cost-to-charge ratios for Florida also enable a sub-analysis using patient costs as a dependent variable for this state.

Regression analyses are performed to test whether hospital and/or surgeon procedure volume for the Whipple procedure is associated with lower inpatient mortality or lower costs per patient. Our results are consistent with previous studies in the medical literature which have identified a significant relation between increased hospital volume and reduced inpatient and three-year mortality for patients undergoing Whipple surgery (J. D. Birkmeyer et al. 1999b;J. D. Birkmeyer et al. 1999a;T. A. Gordon et al. 1998;R. E. Glasgow, S. J. Mulvihill 1996). We go beyond the previous literature by identifying a significant relation between increased surgeon volume and reduced inhospital death, which has not been identified in previous studies of Whipple surgery (E. B. Saettler, W. J. Temple 2000;M. D. Lieberman et al. 1995). We also advance the literature by including in the regressions a Herfindahl index for the Whipple procedure in the hospital's local market, to test whether increased market concentration (lack of competition) has a detrimental effect on patient outcomes or costs. The number of Whipple procedures performed at nearby hospitals is also included as a regressor to test for the potential effects of spillover learning. Year fixed effects are included in the regressions to measure the impact of technological progress in Whipple surgery on inpatient mortality and costs.

The results suggest that both increased hospital volume and surgeon volume are associated with lower inpatient mortality rates and lower charges and costs per patient. Independent time effects are not associated with reduced mortality or lower costs. We also find no evidence that increased local market concentration for the Whipple procedure influences inpatient mortality or average cost per patient. However, the results suggest that spillover learning from nearby hospitals may lead to lower charges per patient.

The findings have important implications for policy recommendations regarding the diffusion of new technologies. If outcome improvements and cost reductions are attributable to increased procedure volume, then current recommendations for minimum volume standards and regionalization of complex technologies should be supported. However, if improvements over time occur for all providers regardless of size, or competition among providers leads to improved outcomes or lower costs, then policy recommendations to regionalize care should be revisited.

The results in this paper suggest that better outcomes and lower charges for the Whipple procedure are due to increased procedure volume. Regionalization can lead to improved outcomes and lower costs, while hospital competition does not. The conclusions are different from those reached for other medical conditions, suggesting that regionalization should be considered on a disease-by-disease basis.

BACKGROUND, DATA AND DESCRIPTIVE STATISTICS

The Whipple procedure is a surgical procedure for patients with localized pancreatic cancer, extra-hepatic bile duct cancer, or cancer of the small intestine. The operation is complex, requiring removal of the head of the pancreas, part of the small intestine, and some of the tissues around it. Although the inpatient mortality rate for the Whipple procedure is relatively high, it is considered to be the most effective method for treating early stage pancreatic cancer. Many patients are willing to accept this surgical risk, because 5-year survival rates for early stage pancreatic cancer are low.[1]

The data for this study come from hospital discharge abstracts provided by the Florida Agency for Health Care Administration for the years 1988 to 1999, the New Jersey Department of Health and Senior Services for 1988 to 1998, and the New York Statewide Planning and Research Cooperative System (SPARCS) for 1989 to 1999. Following previous studies in the

[1]The overall survival rate for pancreatic cancer is 4%. The 5-year survival rate for patients undergoing complete surgical resection has been estimated to be 18% to 24%.

literature, we extracted information on all patients with an ICD-9-CM procedure code of 52.7 (radical pancreaticoduodenectomy) (T. A. Gordon et al. 1998;J. D. Birkmeyer et al. 1999b).The hospital discharge abstracts report total charges for each admission. These charges are adjusted to reflect real dollars using a GDP deflator indexed to 1996. Hospital financial data for Florida is available for all sample years and is used to construct cost-to-charge ratios for each facility. These ratios are used to adjust patient charges to reflect costs in a sub-analysis of Florida patients.

The dataset contains information on 8,145 Whipple procedures performed between the years 1988 and 1999; 3,183 from Florida, 960 from New Jersey, and 4,002 from New York. A total of 432 hospitals and 1,832 surgeons performed the Whipple procedure at least once during the sample period. In 1989 the median hospital performed 3 Whipple surgeries per year; this figure rose to 8 procedures per year by 1998. The largest number of Whipple procedures performed at a hospital in 1989 was 52; by 1998 this figure rose to 132. From 1988 to 1992 the median surgeon performed 1 Whipple procedure per year. This figure increased to 2 per year in 1993 and 3 per year in 1998. The highest volume surgeon in 1989 performed 21 Whipple procedures, and the highest volume surgeon in 1998 performed 40 procedures.

Table 1 provides descriptive statistics on mean characteristics and charges for patients who received the Whipple procedure in 1989, 1994, and 1998. Inpatient mortality fell from 13.7% in 1989 to 9.0% in 1994, but then remained relatively constant through 1998. The casemix severity of patients undergoing Whipple surgery increased slightly over time. Between 1989 and 1998 the average age of the patient population rose from 63.8 years to 65.2 years. The Charlson comorbidity index, a measure of illness severity based on the range of diagnoses in the discharge abstract (P. S. Romano et al. 1993), increased from 2.6 to 3.3 in this same time period.

The reduction in inpatient mortality in spite of a more severely ill patient population suggests that either technological advances or learning in the performance of the Whipple procedure occurred over time.

The average charge per admission for patients receiving Whipple surgery rose from $35,546 in 1989 to $70,289 in 1998. These cost increases may be attributable to the increasing casemix severity of the patient population, increases in input prices, or increases in the quality of the technology used to perform the Whipple procedure.

In general, inpatient mortality rates and average real charges per patient decline monotonically with increasing hospital volume and surgeon volume in each year. Many researchers believe that this volume-outcome relation reflects “practice makes perfect” or “learning by doing” (H. S. Luft et al. 1979;R. Sturm 1999). Providers who perform more surgeries gain experience which leads to improved future outcomes or lower costs. In the context of Whipple surgery, more experienced surgeons may have improved dexterity, or achieve shorter operating times that lower blood loss during surgery. Learning may enable hospitals to develop routines for preventing or treating life threatening complications after surgery. The association between higher volume and lower patient charges may also reflect economies of scale. Even if no learning by doing has occurred, average patient charges may fall if the fixed costs of surgery can be spread over a larger number of patients.

The correlation between annual surgeon volume and hospital volume in the sample is equal to 0.66, indicating that high-volume surgeons tend to operate at high-volume hospitals. Therefore, the descriptive statistics do not allow one to independently identify the effect of surgeon versus hospital procedure volume on inpatient death rates or charges. In addition, Table 1 suggests that higher volume hospitals tend to operate on younger patients. Thus, part of the observed lower mortality rate for high-volume hospitals may be attributable to their propensity to operate on patients who are more able to survive this aggressive procedure. In addition, we wish to test the hypothesis that increased local market competition among hospitals which perform the Whipple procedure may affect outcomes or costs. A multivariate regression is required to examine these issues in more detail.

3. MODEL SPECIFICATION

We estimate the following regression specification to identify the independent effect of surgeon and provider volume and other factors on inpatient mortality:

$$(1) \quad Died_{ihsrt} = f(AnnVol_{ht}, AnnVol_{st}, AnnVol_{rt}, Year_t, Casemix_{it}, \theta_h)$$

where $Died_{ihsrt}$ =1 if patient i treated in hospital h by surgeon s in region r in year t died in hospital, and 0 otherwise. $AnnVol_{ht}$ is the number of Whipple procedures performed in year t in hospital h, and $AnnVol_{st}$ represents procedure volume at the surgeon level. As mentioned previously, these annual volume measures have been hypothesized to capture a learning by doing effect.

Several economists have used cumulative production instead of annual production to estimate learning effects (A. M. Spence 1981;R. Sturm 1999). For example, one could argue that for a patient being treated in 1996, the total number of Whipple procedures ever performed by the operating surgeon prior to 1996 is a more accurate measure of learning. We initially estimated equation (1) including both annual and cumulative measures of procedure volume. However, the correlation between annual and cumulative hospital volume is 0.93; and the correlation between annual and cumulative surgeon volume is 0.85. This high correlation suggests that high-volume providers tend to remain large throughout the sample period, and low-volume providers tend to remain small. Preliminary estimates of equation (1) including both cumulative and annual measures of volume led to unstable regression estimates. This same pattern of multicollinearity was encountered in a previous study of the volume-outcome relation for coronary angioplasty (Ho V. 2002). Therefore, we estimate equation (1) with only annual measures of procedure volume in this study.[2]

Although many researchers would interpret a negative coefficient on the provider volume variables in equation (1) as a learning effect, we cannot rule out the possibility that the relation reflects organizational scale effects. The volume-outcome effect may not be due to learning; instead, it may be attributable to the specialized staff, facilities, and equipment which are more likely to be present at high-volume centers (T. A. Gordon et al. 1999). The multicollinearity between annual and cumulative measures of Whipple procedure volume prevents us from separately identifying the effects of learning by doing versus organizational scale.

Annual procedure volume for all other hospitals in the same region is included as a potential proxy for spillover learning in the performance of the Whipple procedure. Past studies of learning in the economics literature suggest that firms can learn from the production experience of other firms in the same market (A. M. Spence 1981;D. A. Irwin, P. J. Klenow 1994). Regions are defined based on hospital referral regions (HRRs) as defined in the Dartmouth Atlas (Center for the Evaluative Clinical Sciences 2000).[3]

---

[2]Alternative solutions to the multicollinearity problem are not satisfactory or feasible for this study. A ridge regression estimator would yield biased estimates of the learning effect. A principal components specification based on linear combinations of the original explanatory variables would not yield coefficients which are readily interpretable as learning (Greene W.H. 2000). An alternative measure of learning by doing which is uncorrelated with annual procedure volume is necessary to resolve this issue (A. S. Goldberger 1991;J. M. Wooldridge 2000). We have been unable to identify an alternative measure and reserve this issue for future research.

[3]Hospital referral regions represent regional health care markets for tertiary medical care. These regions are defined by aggregating zip codes, then local health care markets where Medicare patients receive most of their care from hospitals within that area. There are 306 hospital referral regions in the U.S. ranging in population in 1996 from 126,329 to 9,288,694.

A recent study found that increased hospital competition led to improved outcomes and lower costs for heart attack patients treated in California (D. P. Kessler, M. B. McClellan 2000). Another study of California patients finds that increased hospital competition reduced inpatient mortality for pneumonia and heart attack patients covered by HMOs, but increases mortality for Medicare patients (G. Gowrisankaran, R. Town 2002). Equation (1) will test the hypothesis the competition influences outcomes for the Whipple procedure by examining the association between local market concentration and mortality rates. Market concentration is measured using a Herfindahl Index based on each hospital's share of patients undergoing the Whipple procedure in each HRR.

Year effects are included in equation (1) to test for potential learning by watching and technological change. Learning by watching represents improvements in production quality or efficiency associated with time, rather than the quantity of production which occurred in the past. This knowledge transfer may occur through the sharing of information at academic conferences, in medical journals, or the transfer of experienced personnel across facilities. Learning by watching may also occur if experience leads to the creation of new surgical technologies which diffuse quickly amongst providers.

Patient-specific casemix variables are included to control for disease severity. These variables include patient age, gender, indicator variables for the 8 conditions that comprise the Charlson comorbidity index,[4] and dummy variables distinguishing the 5 indications for Whipple

[4]These comorbidities are: a previous myocardial infarction, peripheral vascular disease, chronic pulmonary disease, mild liver disease, mild to moderate diabetes, diabetes with chronic complications, renal disease, and moderate or severe liver disease.

surgery.[5] Length of stay is included in the inpatient mortality equation, because one is more likely to observe in-hospital death for patients who have longer hospitalizations.[6] Indicator variables for whether or not the patient was admitted to a teaching hospital, as well as dummy variables for Florida and New Jersey patients were included in equation (1). Patient mortality is also hypothesized to depend on a vector of unobserved hospital characteristics θ_h.

Equation (1) is estimated including linear terms for each of the procedure volume variables. Quadratic terms for each of the volume variables are included in the reported specification if they were precisely estimated in preliminary estimates. Preliminary estimates also included interactions of provider volume with a state dummy variable. These interaction terms were included in the reported specification if they were precisely estimated.

Equation (1) is estimated using the xtlogit command in Stata 7.0. The xtlogit model allows one to estimate a panel data logit model with either conditional fixed effects or random effects to account for hospital-level unobserved heterogeneity. A Hausman test is used to select between the random effects and conditional fixed effects specifications (J. Hausman 1978). The random effects specification is reported if the null hypothesis that it yields both consistent and efficient estimates cannot be rejected.

A similar regression specification characterizes the relation between individual patient charges and provider volume:

$$(2) \quad \log(Charge)_{ihsit} = f(AnnVol_{ht}, AnnVol_{st}, AnnVol_{it}, Year_t, Casemix_{it}, \theta_h)$$

[5]The 5 indications for surgery are: pancreatic cancer, extra-hepatic bile duct cancer, duodenal cancer, benign pancreatic disease, and other diagnoses. Patients receiving the Whipple surgery due to trauma are excluded from the sample.

[6]We also estimated the inpatient mortality regression excluding length of stay as an explanatory variable. The coefficients on all of the volume variables remained the same to the second decimal point. These results are available from the authors upon request.

The dependent variable is the natural log of charges for the patient's stay in hospital. The explanatory variables are the same as those included in the inpatient mortality regression, although length of stay is not included as a regressor. Regression diagnostics suggested that the residuals from the log(Charge) regressions displayed kurtosis which was too great to justify estimation with generalized linear models (W. G. Manning, J. Mullahy 2001). Therefore, equation (2) is estimated using OLS with hospital-specific fixed effects and the cluster option in Stata 7.0. The cluster option is utilized to account for potential heteroskedasticity in the error term across hospitals. Because the hospital-specific fixed effects span the indicator variable for hospital teaching status and the state dummy variables, these variables are excluded from equation (1).

We estimate equation (2) for all three states using charges as the dependent variable, because we do not have data on costs for New Jersey and New York. Although increased costs are likely to lead to increased charges, the two variables are not perfectly collinear. Hospitals with strong market power may raise charges in an attempt to increase revenue, even if their operating costs have not changed. Therefore, equation (2) is also estimated for Florida patients, with the available data on hospital cost for each patient used as the dependent variable instead of charges. This sub-analysis allows us to examine whether volume and year effects persist when one focuses on costs.

The Florida patient discharge database did not begin to record surgeon identifiers until 1992. Therefore, 731 patients from Florida in the years 1988 to 1991 are not included in the regression sample. In addition, 29 patients in other years from Florida, 33 New Jersey patients, and 25 New York patients were missing surgeon identifiers in other years and are excluded from the regressions.

4. RESULTS

4a. Inpatient Mortality Results

Column 1 of Table 2 provides logit estimates of the determinants of inpatient mortality for Whipple surgery patients. A Hausman test indicated that we could not reject the null hypothesis that the random effects specification is consistent and efficient ($\chi^2(33)=39.01$, $p=0.22$). Therefore, random effects estimates are reported in Column 1.

The parameter estimates indicate that procedure volume at both the hospital and the surgeon level leads to lower probabilities of inpatient mortality. However, increases in surgeon volume lead to reduced mortality at a decreasing rate. These results are precisely estimated and consistent with the hypothesis that learning by doing or organizational scale economies reduce inpatient mortality for the Whipple procedure.

The coefficients on annual region volume and its square suggest that increases in procedure volume by other hospitals in the same hospital referral region increase the probability of inpatient mortality at a given hospital, although at a decreasing rate. This result is puzzling and difficult to interpret. This issue can be explored in more detail if one conducts simulations at specific levels of hospital referral region volume to compare the relative effects of the linear and quadratic term on expected inpatient mortality rates.

We used the estimates in Column 1 to conduct simulations to compare the relative magnitude of each of the volume effects on inhospital death. We used the characteristics of each patient in the sample to predict their probability of death at particular volume levels. For example, the first row in Table 3 indicates that if all patients in the sample had been treated in a hospital where they were the only Whipple patient that year, and the number of other Whipple procedures in the region were equal to 50, then the expected mortality rate would be 10.6%. Thus, the predictions are calculated using fixed values of hospital, surgeon, and region volume; but allowing all other explanatory variables to take on their actual values in the sample.

The predictions in Table 3 are calculated at procedure volumes which are representative of the majority of the sample. For example, in 1998 a hospital at the 75th percentile of the volume distribution operated on 5 patients, a hospital at the 95th percentile performed 15 procedures, and the largest hospital treated 132 patients. Therefore, we chose to predict the association between hospital volume and inpatient mortality up to 20 procedures per year. The predictions suggest that if a hospital performs only one Whipple procedure per year (so that surgeon volume is also equal to one), then the expected mortality rate for the patient is 10.6%. However, if a surgeon who only performs one Whipple procedure per year conducts the operation in a hospital which performs 20 Whipple procedures, the expected inpatient death rate falls to 8.3 percent. Being treated by a high-volume surgeon reduces inpatient death rates even more. A patient treated by a surgeon who performs only one Whipple procedure per year in a hospital that performs five procedures has an expected inpatient mortality rate of 10.1%. However, if only one surgeon performs all the procedures in a given hospital that treats five patients per year, then the patient's expected inhospital mortality rate falls to 7.7%. In fact, the expected inpatient mortality rate for a surgeon performing one Whipple procedure per year is more than twice as large as the expected rate for a surgeon performing 20 procedures (10.6% versus 4.8%).

We also predicted inpatient mortality rates at varying levels of procedure volume for other hospitals in the same referral region, fixing hospital volume constant at 10 procedures, and surgeon volume at five procedures. As the number of Whipple procedures at other hospitals rises from 10 per year to 100, inpatient mortality rises from 6.3% to 7.8%. However, if the number of Whipple surgeries performed by nearby hospitals increases to 200 per year, expected inpatient mortality falls to 6.4%. In 1998, the 25th, 50th, 75th, and 95th percentiles of the distribution for procedure volume at other hospitals is equal to 6, 15, 26, and 140 procedures per year respectively, and the highest value of the distribution is equal to 286. Therefore, at very high levels of procedure volume for nearby hospitals in the sample, the estimates suggest that spillover learning may exist. However, for most patients in the sample, increases in nearby hospitals' experience seems to raise inpatient mortality. As the propensity to perform the Whipple procedure increases in a certain region, patients at greater risk of death may be more likely to undergo surgery. We have attempted to control for differences in patient severity using information on age, comorbidities, and the indication for surgery. However, the regional volume variable may be capturing unobserved increased patient severity in high volume regions.

Returning to the regression estimates in Table 2, note that the coefficient on the Herfindahl index in Column (1) is imprecisely estimated. Therefore, we find no evidence that increases in market competition in the hospital referral region lead to reduced inpatient mortality. The coefficients on all of the year dummy variables are negative (relative to patients admitted in 1988 or 1989). However, many of the coefficients are imprecisely estimated. In fact, computation of a Wald statistic suggests that we cannot reject the null hypothesis that the coefficients on the year effects are jointly equal to 0 ($\chi^2(10)=9.47$, $p=0.49$). Therefore, we find no definitive evidence of reductions in inpatient mortality attributable to technological progress or learning by watching for the Whipple procedure.

4b. Charge Estimates

Column 2 of Table 2 provides estimates of the determinants of the natural log of per patient charges for the Whipple procedure. Patients with missing surgeon identifiers or charge information were not in the analysis, reducing the sample size to 7,114 patients. A Hausman test rejected the null hypothesis that the random effects specification is consistent and efficient [χ^2=164.78, p<0.001]. Therefore, the estimates including hospital-specific fixed effects are reported.

The estimates in Column 2 indicate that increased hospital procedure volume is associated with lower charges per patient, although the effect on charges occurs at a decreasing rate. In addition, higher surgeon volume for the Whipple procedure is associated with lower charges per patient. Therefore, learning by doing and/or economies of scale lead to reduced charges per patient.

Increases in the number of Whipple procedures performed by nearby hospitals also appears to lower charges per patient. These results are consistent with the hypothesis that spillover learning between hospitals may tend to lower the costs of performing a procedure. The Herfindahl index in Column 2 is imprecisely estimated. Therefore, we find no evidence that increased competition between hospitals leads to reduced charges. One may be concerned that the number of other Whipple procedures performed in the hospital referral region is a proxy for increased competition, which may explain the imprecise estimate of the Herfindahl coefficient. We re-estimated the charges equation excluding hospital referral region volume, and the coefficient on the Herfindahl index remained imprecisely estimated (t=-0.24). The year effects are precisely estimated and indicate increasing charges throughout the sample period. Therefore, we find no evidence that learning by watching or technological change has helped to reduce the costs of performing the Whipple surgery over time.

Column 3 of Table 2 reports estimates of the determinants of costs for Whipple surgery patients in Florida, where cost-to-charge ratios for each hospital over time are available. Both increased hospital and surgeon volume are associated with reduced costs. However, the effect of increased Whipple procedure volume by nearby hospitals becomes imprecisely estimated. The Herfindahl index also remains imprecisely estimated. We are missing year effects for 1990 through 1992, because surgeon identifiers are unavailable for these years in Florida. Relative to 1992, the year effects suggest no tangible difference in costs per patient. These results suggest that the increase in charges observed over time in Column 2 are not attributable to increasing costs, but instead due to increases in hospital cost-to-charge ratios over time. Again, the results find no evidence of reduction in costs attributable to technological change or learning by watching. Any cost reductions are attributable to learning by doing or economies of scale.

We can use the estimates in Columns 2 and 3 to estimate the cost of treating patients in our sample in hospitals, regions, and by surgeons with particular procedure volumes. Predicted charges and costs are calculated using the smearing estimator to adjust for potential prediction error due to non-normality of the error term in the regression equation (N. Duan 1983;W. G. Manning 1998). The results appear in Table 3. Column 5 of Table 3 indicates that if all patients in the sample had been treated in a hospital where they were the only Whipple patient that year, and the number of other Whipple procedures in the region were equal to 50, then the average expected costs per patient would be $72,060. If the surgeon continues to perform only one procedure per year, but does so in a hospital operating on 20 patients per year, then expected patient costs fall to $58,364. Similar reductions in costs are noted for increases in surgeon volume. A patient treated by a surgeon who performs only one Whipple procedure per year in a hospital that performs 10 procedures has an expected charge of $64,938. However, if one surgeon performs all the procedures in a hospital that treats 10 patients per year, the expected charge drops to $59,495.

The cost savings associated with spillover learning also appear to be substantial. Suppose one predicts expected charges for a patient who is treated by a surgeon who performs 5 procedures in a given year in a hospital which treats a total of 10 patients. As the number of Whipple procedures at other hospitals rises from 10 per year to 200, expected charges fall from $68,172 to $44,985 per patient.

Column 6 of Table 3 reports cost predictions based on the regression estimates of the determinants of costs for patients in Florida. These figures also suggest substantial reductions in average cost per patient associated with increases in hospital and surgeon volume. These results suggest that regionalization of the Whipple procedure may lead to substantial cost savings.

A brief examination of the coefficients on the patient characteristic variables in the mortality and charge equations indicates that the estimates are consistent with the clinical prognosis for these patients. Relative to patients under age 60, older patients have a higher probability of death in hospital and greater charges. Women are less likely to die in hospital than men; they also have lower charges than men. Comorbidities tend to increase the probability of inpatient mortality and patient charges. The estimates suggest that presence of diabetes reduce the probability of death in hospital. This result has been identified in several studies using administrative data in the past. The effect is attributed to a lower propensity of coding these conditions for patients at increased risk of death who face multiple highly severe complications while hospitalized (L. I. Iezzoni 1992;S. F. Jencks et al. 1988). Finally, patients treated in teaching hospitals have a substantially lower probability of death than patients treated in non-teaching facilities. The marginal effect implied by the estimated coefficient in the logit regression suggests that the expected death rate for patients treated in teaching hospitals is lower by 1.7 percentage points.

4c. Alternative Specifications

We estimated a variety of alternative specifications to examine the robustness of the results. The quadratic specification for some of the volume measures suggests that the volume-outcome effect "wears off" at higher levels of procedure volume. Given that some recommend regionalization of the Whipple procedure, it would be helpful to determine a cutoff point beyond which volume increases fail to provide meaningful reductions in mortality. We took a closer look at this issue by estimating the determinants of inpatient mortality using categorical dummy variables for procedure volume. We divided the hospital, surgeon, and region volume measures into approximate quartiles, and first tested for significant differences in inpatient mortality relative to the lowest volume quartile. For both hospital and surgeon volume, we found that the coefficient on the highest volume quartile was negative and precisely estimated.[7] We then divided the highest quartile for both hospital and surgeon volume in half and tested to see whether the highest eighth of the hospital or surgeon volume distribution had lower inpatient mortality rates than the second-highest eighth. In the logit regression of mortality, the coefficient for the highest eighth of the hospital distribution relative to the second-highest eighth is equal to -0.448 (t=-1.42); and the same coefficient for the surgeon distribution is equal to -0.351 (t=-1.26). Therefore, the probability of inpatient death continues to decrease for the highest eighth of both the hospital and physician volume distributions; however, neither effect is precisely estimated. These imprecise estimates may be due to the relatively limited sample size

[7]These results are available from the author upon request. Not all of the other volume quartiles were precisely estimated. However, we do not discuss these results in detail, because we are primarily interested in examining the volume-outcome effect at the highest levels of procedure volume.

of 7,327 patients in the inpatient mortality regressions. That is, the sample size may be insufficient to detect a clinically meaningful difference in mortality rates between the top two eighths of the hospital and physician volume distributions. A larger sample would aid in identifying the point at which higher procedure volumes cease to provide additional mortality reductions.

Another alternative specification included an interaction term between hospital and surgeon Whipple procedure volume along with the variables specified in the inpatient mortality regression in Table 2. This coefficient was equal to 0.00067, but imprecisely estimated (t=1.54). The estimate provides weak evidence that the beneficial effect of being treated by a high-volume surgeon at a high-volume hospital may be slightly smaller than implied by the combination of the independent effects of hospital and surgeon volume reported in Tables 2 and 3. A similar interaction term was added to the patient charge regression specified in Column 2 of Table 2. This estimate was equal to 0.00004, but also imprecisely estimated.

We also experimented with including interaction effects between the hospital and surgeon volume measures and time period. Time periods were defined by splitting the sample into three categories (1988-1991, 1992-1995, and 1996-1999) We examined these interaction terms to test whether the relationship between procedure volume and inpatient mortality became flatter or steeper over time. None of these interaction terms were precisely estimated.

5. CONCLUSIONS

Past studies identifying a cross-sectional association between higher procedure volume and lower inpatient mortality have been used to recommend minimum volume standards or regionalization of complex surgeries. These minimum volume standards have been incorporated into state Certificate of Need regulations for a number of procedures. In addition, The Leapfrog Group, a large coalition of major U.S. employers and health care purchasers, is encouraging employees and customers to select high quality hospitals for care; where quality is measured in part by the volume of procedures performed each year.

Yet cross-sectional studies of the volume-outcome relation cannot take into account the possibility that outcomes may improve over time for all providers, regardless of procedure volume. For instance, a recent study of patients receiving coronary angioplasty found improvements in outcomes associated with increases in procedure volume over the period 1984 to 1996 were quite small (Ho V. 2002). Reductions in inpatient mortality associated with learning by watching and/or technological change were substantially larger. In addition, analysis of heart attack patients in California found that increased competition leads to reduced mortality and lower costs of care (D. P. Kessler, M. B. McClellan 2000). Similar results have been found for HMO patients admitted to hospitals for pneumonia (G. Gowrisankaran, R. Town 2002). Thus, given that regionalization could lead to reduced hospital competition and reduced access to advanced surgeries in outlying areas, this policy option should be exercised with caution.

This paper builds on the previous literature by comparing the relative importance of provider-specific volume versus hospital competition, learning by watching, and spillover effects in explaining outcomes and charges for a complex and expensive cancer surgery. Provider effects at both the hospital and surgeon level are assessed. The results indicate that both increased hospital volume and surgeon volume are associated with lower inpatient mortality rates and lower charges per patient for Whipple surgery. Cumulative procedure volume and annual procedure volume are highly correlated, so that one cannot distinguish between learning by doing versus scale effects. The results do not suggest that increased hospital competition or learning by watching/technological change improves outcomes or reduces charges. However, we do find that spillover learning across hospitals leads to reduced charges per patient. Nevertheless, for the Whipple procedure, regionalization would lead to improved outcomes and lower costs, while hospital competition does not.

The results in this study are in direct contrast to findings regarding the high degree of learning by watching/technological change found for coronary angioplasty and the benefits of competition identified in the market for care of heart attack patients. Yet these differences in findings are consistent with the clinical and market factors characterizing each patient group. The Whipple procedure is one of the most complex operations a surgeon can perform, and surgery can lead to several life threatening complications. Therefore, it is reasonable to believe that outcomes depend critically upon experience accumulated by both the surgeon and critical care staff at the hospital. In contrast, coronary angioplasty requires experience; however, outcomes have improved substantially over time due to improved catheters, stents, and drugs associated with the procedure (D. M. Cutler, M. B. McClellan 2001).

The market for the Whipple procedure is also quite small. Most hospitals will perform only 1 or 2 Whipple procedures per year, while they often care for hundreds of heart attack patients. Therefore, hospitals are more likely to compete for coronary care patients than they are to compete for patients with pancreatic cancer. These results suggest that efforts to regionalize complex surgeries should be considered on a disease-by-disease basis. The decision to regionalize should be based upon the magnitude of the identified volume-outcome relation, improvements in outcomes associated with learning by watching/technological change, the potential impact of regionalization on access to care, and the potential for competition to improve outcomes and costs.

The analysis in this paper still contains a number of limitations. Cumulative and annual procedure volume are highly correlated, so that one cannot distinguish between learning by doing and scale effects at the hospital and physician level. This same collinearity was found in a previous longitudinal study of patients receiving coronary angioplasty (Ho V. 2002). This pattern suggests that hospitals which are big tend to stay big; and hospitals which are small tend to stay small. Identification of natural experiments which lead to marked changes in annual volume over a short period of time may be necessary to disentangle the effects of learning by doing versus scale effects for surgery outcomes.

The analysis in this paper does not directly address the hypothesis that the relation between procedure volume and outcomes is endogenous. Learning may lead to improved outcomes and lower costs; but higher quality patient care may also attract more patients (H. S. Luft et al. 1979). However, past studies found no evidence that patients avoid hospitals with well-publicized poor outcomes or favor hospitals with good ones (M. R. Chassin et al. 1996;E. C. Schneider, A. M. Epstein 1998). Nevertheless, fixed or random effects were included in all specifications to control for systematic unobserved differences in quality across hospitals. Duplicating the analysis with instrumental variables would provide additional information on the impact of procedure volume on outcomes and costs. Conducting instrumental variables analysis requires identification of variables that influence provider volume, but not outcomes or costs. More studies should be conducted to identify the factors which lead some providers to be high-volume and others to be low-volume.

Finally, the analysis of market competition in this study is limited. Local markets are defined based on hospital referral regions, which may be endogenous; increased competition may lead hospitals to draw patients from a larger geographic area, which increases the size and therefore the number of patients in the hospital referral region. Recent studies have constructed exogenous measures of hospital competition using distance from the nearest hospital as an instrumental variable for market concentration (D. P. Kessler, M. B. McClellan 2000;G. Gowrisankaran, R. Town 2002). Duplicating this computationally expensive methodology is beyond the scope of this paper, because we have data on three states; in contrast, the previous studies were combined to California. The absence of competitive effects should be re-examined with alternate definitions of market concentration in future studies.

Table 1: Trends in Patient Characteristics and Charges by Whipple Volume (selected years)

Year		In-hospital death (%)		Mean Charge per Patient		Mean Age		Charlson Index	
1989, N=518		13.7		$35546		63.8		2.6	
Hosp. Vol.		Hosp	Surg.	Hosp.	Surg.	Hosp.	Surg.	Hosp.	Surg.
1-3		18.3		29543		64.7		2.4	
4-9		10.6		32690		63.5		2.8	
10+	Surg. Vol	4.5		31608		61.4		3.0	
	1		20.9		23157		62.7		2.2
	2-4		8.1		22709		64.2		2.8
	5+		4.8		22918		61.6		3.0
1994, N=720		9.0		61029		64.7		3.3	
Hosp. Vol.		Hosp	Surg.	Hosp.	Surg.	Hosp.	Surg.	Hosp.	Surg.
1-3		13.8		65515		65.5		3.2	
4-9		7.9		55790		64.8		3.1	
10+	Surg. Vol.	4.3		60172		63.5		3.7	
	1		11.0		65582		64.5		3.3
	2-4		11.7		63497		65.5		3.1
	5+		3.3		49805		64.1		3.7
1998, N=1062		8.9		70289		65.2		3.3	
Hosp. Vol.		Hosp	Surg.	Hosp.	Surg.	Hosp.	Surg.	Hosp.	Surg.
1-3		16.8		83825		66.3		3.1	
4-9		10.9		73098		66.9		3.5	
10+	Surg. Vol.	3.6		61890		63.5		3.3	
	1		11.9		82902		65.1		3.4
	2-4		11.9		77554		66.0		3.1
	5+		4.8		57018		64.7		3.4

Table 2: Determinants of Mortality, Charges, and Costs

	(1) Logit (Mortality)		(2) Ln (Charges)		(3) Ln (Costs)	
Annual Hospital Volume	-0.014	(-3.44)	-0.012	(-2.51)	-0.023	(-3.81)
Annual Hospital Volume* NJ	-0.173	(-2.74)				
Annual Hospital Volume* FL			-0.015	(-2.07)		
Annual Hospital Volume^2			-4.75E-05	(1.83)	-2.3E-04	(4.39)
Annual Hospital Volume^2* FL			-1.9E-04	(3.11)		
Annual Surgeon Volume	-0.084	(-3.10)	-0.010	(-6.51)	-0.007	(-2.31)
Annual Surgeon Volume^2	0.002	(1.93)				
Annual Region Volume	0.005	(1.64)	-0.002	(-2.86)	-0.001	(-0.41)
Annual Region Volume^2	-2.45E-05	(-1.87)				
Herfindahl Index	-0.090	(-0.27)	-0.007	(-0.07)	-0.219	(-1.24)
1990	-0.575	(-2.22)	0.065	(0.77)		
1991	-0.408	(-1.69)	0.273	(4.97)		
1992	-0.229	(-1.06)	0.459	(8.73)		
1993	-0.190	(-0.90)	0.610	(11.45)	-0.009	(-0.18)
1994	-0.367	(-1.70)	0.657	(12.36)	-0.046	(-0.85)
1995	-0.381	(-1.75)	0.646	(9.90)	-0.015	(-0.26)
1996	-0.256	(-1.19)	0.662	(9.59)	-0.040	(-0.86)
1997	-0.409	(-1.90)	0.799	(12.59)	-0.030	(-0.60)
1998	-0.192	(-0.90)	0.980	(14.73)	0.035	(0.52)
1999	-0.395	(-1.72)	0.954	(14.04)	0.046	(0.89)
Age 60-69	0.427	(3.21)	0.098	(4.54)	0.087	(2.39)
Age 70-79	0.946	(7.49)	0.158	(7.95)	0.132	(3.52)
Age 80	1.487	(8.91)	0.216	(7.38)	0.168	(3.22)
Female	-0.275	(-3.10)	-0.050	(-2.68)	-0.053	(-2.41)
Myocardial Infarction	-0.361	(-0.88)	-0.109	(-2.07)	-0.120	(-1.73)
Peripheral vascular disease	0.264	(0.90)	0.028	(0.57)	-0.010	(-0.14)
Chronic pulmonary disease	0.072	(0.50)	0.045	(1.72)	0.058	(1.56)
Mild liver disease	0.920	(3.01)	0.114	(1.96)	0.153	(1.29)
Mild/moderate diabetes	-0.476	(-3.43)	-0.052	(-2.78)	-0.014	(-0.51)
Diabetes with chronic complications	-0.881	(-2.79)	-0.026	(-0.63)	0.041	(0.64)
Renal disease	3.249	(9.87)	0.596	(8.03)	0.533	(4.93)
Moderate/severe liver disease	1.951	(5.37)	0.232	(2.54)	0.341	(2.16)
Extra-hepatic bile duct cancer	-0.239	(-2.11)	0.031	(1.58)	0.014	(0.43)
Duodenal cancer	0.006	(0.03)	0.128	(3.85)	0.146	(2.62)
Benign pancreatic disease	-0.888	(-3.50)	-0.004	(-0.14)	-0.048	(-1.24)
Other indication for Whipple	-0.050	(-0.34)	0.122	(3.93)	0.070	(1.46)
Length of stay	0.004	(2.14)				
Teaching hospital	-0.276	(-2.02)				
FL	-0.149	(-0.97)				
NJ	0.602	(2.42)				
Constant	-2.025	(-7.37)	10.36	(130.92)	10.435	(129.78)
	random effects		fixed effects		fixed effects	
N	7327		7114		2423	

t-statistics in parentheses

Table 3: Predicted Mortality Rates, Charges, and Costs

(1) Hospital Volume	(2) Surgeon Volume	(3) Region Volume	(4) Predicted Mortality	(5) Predicted Charges	(6) Predicted Cost
1	1	50	10.6	72,060	42,280
2	1	50	10.5	71,205	41,358
3	1	50	10.4	70,367	40,475
4	1	50	10.2	69,545	39,630
5	1	50	10.1	68,739	38,821
10	1	50	9.5	64,938	35,263
20	1	50	8.3	58,364	30,140
2	2	50	9.8	70,516	41,080
3	3	50	9.0	69,011	39,933
4	4	50	8.3	67,545	38,836
5	5	50	7.7	66,116	37,787
10	10	50	5.5	59,495	33,185
20	10	50	4.8	53,472	28,364
4	2	50	9.5	68,872	39,363
10	5	50	7.2	62,460	34,324
10	5	10	6.3	68,172	35,064
10	5	25	6.7	65,971	34,785
10	5	100	7.8	55,987	33,420
10	5	200	6.4	44,985	31684

Reference List

1. Birkmeyer, J. D., Finlayson, S. R. G., Tosteson, A. N. A., Sharp, S. M., Warshaw, A. L., Fisher, E. S. "Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy." *Surgery*, 1999a, *125* (3), 250-256.

2. Birkmeyer, J. D., Siewers, A. E., Finlayson, S. R., Stukel, T. A., Lucas, F. L., Batista, I., Welch, H. G., Wennberg, D. E. "Hospital volume and surgical mortality in the United States." *N Engl J Med*, April 2002, *346* (15), 1128-1137.

3. Birkmeyer, J. D., Warshaw, A. L., Finlayson, S. R. G., Grove, M. R., Tosteson, A. N. A. "Relationship between hospital volume and late survival after pancreaticoduodenectomy." *Surgery*, 1999b, *126* (178), 183.

4. Center for the Evaluative Clinical Sciences. The Dartmouth Atlas of Health Care 1999. http://www.dartmouthatlas.org/atlaslinks/99atlas.php . 2000.
Ref Type: Electronic Citation

5. Chassin, M. R., Hannan, E. L., DeBuono, B. A. "Benefits and hazards of reporting medical outcomes publicly." *N Engl J Med*, 1996, *334* (6), 394-398.

6. Cutler, D. M., McClellan, M. B. "Is technological change in medicine worth it?" *Health Affairs*, 2001, *20* (5), 11-29.

7. Duan, N. "Smearing estimate: a nonparametric retransformation method." *Journal of the American Statistical Association*, 1983, *(78)* 605-610.

8. Glasgow, R. E., Mulvihill, S. J. "Hospital volume influences outcome in patients undergoing pancreatic resection for cancer." *Western Journal of Medicine*, November 1996, *165* (5), 294-300.

9. Goldberger, A. S. *A Course in Econometrics*. Cambridge, MA: Harvard University Press, 1991.

10. Gordon, T. A., Bowman, H. M., Bass E.B., Lillemoe, K. D., Yeo, C. J., Heitmiller, R. F., Choti, M. A., Burleyson, G. P., Hsieh, G., Cameron, J. L. "Complex gastrointestinal surgery: Impact of provider experience on clinical and economic outcomes." *J Am Coll Surg*, 1999, *189* (1), 46-56.

11. Gordon, T. A., Bowman, H. M., Tielsch, J. M., Bass, E. B., Burleyson, G. P., Cameron, J. L. "Statewide regionalization of pancreaticoduodenectomy and its effect on in-hospital mortality." *Annals of Surgery*, 1998, *228* (1), 71-78.

12. Gowrisankaran, G. and Town, R. Competition, payers, and hospital quality. 2002.
Ref Type: Unpublished Work

13. Greene W.H. "The classical multiple linear regression model: specification and estimation," *Econometric Analysis*. Upper Saddle River, NJ: Prentice-Hall, Inc., 2000, 210-270.

14. Hannan, E. L., Kilburn Jr., H., Bernard, H., O'Donnell, J. F., Lukacik, G., Shields, E. P. "Coronary artery bypass surgery: The relationship between inhospital mortality rate and surgical volume after controlling for clinical risk factors." *Medical Care*, 1991, *29* (11), 1094-1107.

15. Hausman, J. "Specification tests in econometrics." *Econometrica*, 1978, *46* 1251-1271.

16. Ho V. "Learning and the evolution of medical technologies: The diffusion of coronary angioplasty." *Journal of Health Economics*, 2002, *21* (5), 873-885.

17. Hughes, R. G., Garnick, D. W., Luft, H. S., McPhee, S. J., Hunt, S. S. "Hospital volume and patient outcomes: The case of hip fracture patients." *Medical Care*, 1988, *26* (11), 1057-1067.

18. Iezzoni, L. I. "Comorbidities, complications, and coding bias: does the number of diagnosis codes matter

in predicting in-hospital mortality?" *The Journal of the American Medical Association* , 1992, *267 (16)* 2197.

19. Irwin, D. A., Klenow, P. J. "Learning-by-doing spillovers in the semiconductor industry." *Journal of Political Economy*, 1994, *102* (6), 1200-1227.

20. Jencks, S. F., Williams, D. K., Kay, T. L. "Assessing hospital-associated deaths from discharge data." *The Journal of the American Medical Association*, 1988, *260* 2240-2246.

21. Kessler, D. P., McClellan, M. B. "Is hospital competition socially wasteful?" *The Quarterly Journal of Economics*, 2000, *115* (2), 577-615.

22. Lieberman, M. D., Kilburn, H., Lindsey, M., Brennan, M. F. "Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy." *Annals of Surgery*, November 1995, *222* (5), 638-645.

23. Luft, H. S., Bunker, J., Enthoven, A. "Should operations be regionalized? An empirical study of the relation between surgical volume and mortality." *The New England Journal of Medicine*, 1979, *301* (25), 1364-1369.

24. Manning, W. G. "The logged dependant variable, heteroskedasticity, and the retransformation problem." *Journal of Health Economics*, 1998, *17* 283-295.

25. Manning, W. G., Mullahy, J. "Estimating log models: to transform or not to transform?" *Journal of Health Economics*, 2001, *20* 461-494.

26. Phillips, K. A., Luft, H. S., Ritchie, J. L. "The association of hospital volumes of percutaneous transluminal coronary angioplasty with adverse outcomes, length of stay, and charges in California." *Medical Care*, 1995, *33* (5), 502-514.

27. Romano, P. S., Roos, L. L., Jollis, J. G. "Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: Differing perspectives." *Journal of Clinical Epidemiology*, 1993, *46* (10), 1075-1079.

28. Saettler, E. B., Temple, W. J. "The surgeon as a prognostic factor in the management of pancreatic cancer." *Surgical Oncology Clinics of North America*, January 2000, *9* (1), 133-142.

29. Schneider, E. C., Epstein, A. M. "Use of public perfromance reports: a survey of patients undergoing cardiac surgery." *The Journal of the American Medical Association*, 1998, *279* (20), 1638-1642.

30. Spence, A. M. "The learning curve and competition." *Bell Journal of Economics*, 1981, *12* (1), 49-70.

31. Sturm, R. "Cost and quality trends under managed care: Is there a learning curve in behavioral health carve-out plans?" *Journal of Health Economics*, 1999, *18* (5), 593-604.

32. Wooldridge, J. M. *Introductory Econometrics: A Modern Approach*. Mason, OH: South-Western College Publishing, 2000.